nature portfolio

https://doi.org/10.1038/s41586-024-08400-1

Supplementary information

The sequence–structure–function relationship of intrinsic ERα disorder

In the format provided by the authors and unedited

Supplementary Methods

Coarse-grained simulation

The CALVADOS2 model[76,77] was employed for coarse-grained molecular dynamics simulations. In this model, each amino acid was represented as a single bead. Simulations were performed using Langevin dynamics with a friction coefficient of 0.01 ps^{-1} and a time step of 10 fs. The temperature was maintained at 310 K, with an ionic strength of 150 mM and a pH of 7.4. For each protein sequence, simulations were conducted for 1 μs. The initial 200 ns of each trajectory was designated as the equilibration period and excluded from subsequent analysis, leaving 800 ns of equilibrated sampling for analysis.

The conformational properties of the simulated proteins were characterized by calculating the Flory exponent (v) using its relationship with the radius of gyration (R_g) in a self-avoiding walk model. This calculation followed the methodology described in Eq. 7 and Eq. 8 of Ref. [78].

Initial simulations of ER-NTD yielded a radius of gyration slightly larger than experimental SEC-SAXS measurements. To address this discrepancy, the model was refined by adjusting the interaction strength (ε) in the pairwise hydrophobic interaction term. The ε parameter was optimized by minimizing the χ^2 value, defined as

$$\chi^2 = \frac{1}{N_q}\sum \frac{(I_{\exp}(q) - I_{sim}(q))^2}{\sigma^2_{\exp}(q)},$$

where $I_{sim}(q)$ and $I_{exp}(q)$ are the simulated and experimental scattering intensities, respectively, $\sigma_{exp}(q)$ is the experimental error of $I_{exp}(q)$, and N_q is the number of scattering vector q data points.

As shown in Extended Data Fig. 10b, the optimal value was determined to be ε = 0.23 kcal/mol, which reproduces SEC-SAXS data of ER-NTD (Extended Data Fig. 10c). This refined parameter was applied to simulations of the 48 nuclear receptor NTDs (Extended Data Fig. 10a).

Multi-Gaussian relaxation analysis

The model introduced by Klein-Seetharaman et al.[24] was employed to fit the NMR relaxation parameters. The fitting equation is as follows:

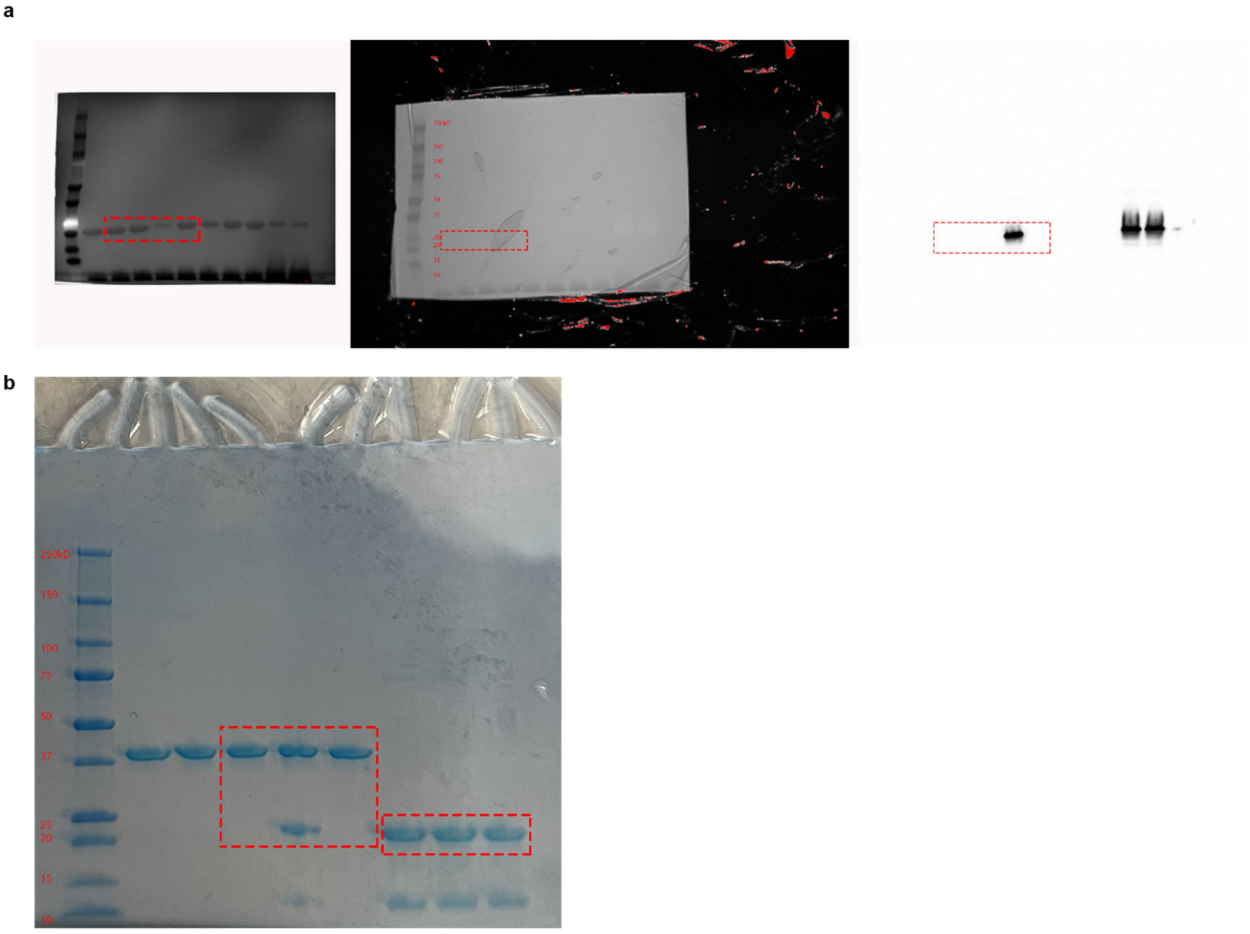
$$R(i) = \frac{R_2(i)}{R_1(i)} = R_0 \sum_{j=1}^{N} \exp\left(-\frac{|i-j|}{\lambda_0}\right) + \sum_{cluster} R_{cluster} \exp\left(-\frac{(i - x_{cluster})^2}{\lambda^2_{cluster}}\right),$$

where R_2/R_1 is the ratio of transverse to longitudinal relaxation rates, i is the amino acid position, N is the residue account, R_0 is the baseline relaxation rate ratio, λ_0 is the persistence length (in number of residues), $x_{cluster}$ is the center position of each cluster, $R_{cluster}$ is the amplitude of each cluster, and $\lambda_{cluster}$ is the width of each cluster. In this equation, the first term characterizes the intrinsic relaxation rates from a random coil, while the second term corresponds to clusters of amino acids with backbone motions deviating from a random coil. The persistence length λ_0 was set to 7 residues, consistent with the original literature[24].

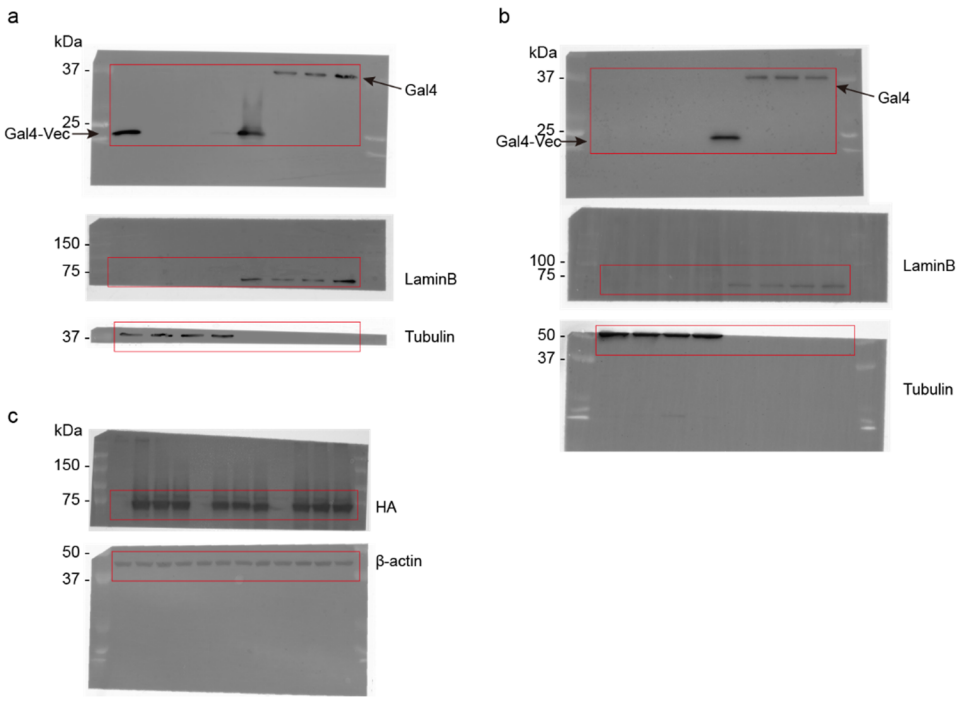
The analysis revealed that two additional clusters were sufficient to fit the data effectively. The cluster location is defined as the Full Width at Half Maximum (FWHM) of the Gaussian function for each cluster. The fitting yielded the following parameters:

cluster name	$x_{cluster}$ (amino acid position)	$R_{cluster}$	$\lambda_{cluster}$	cluster location (amino acid position)
		0.31 (R_0)	7 (λ_0)	
Cluster-I	62	4.75	22	43-81
Cluster-II	130	4.79	26	108-152

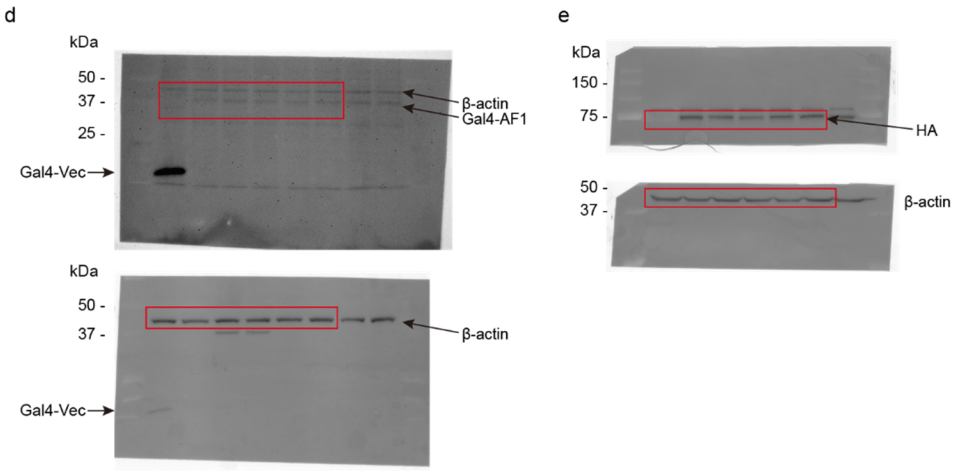
Supplementary Figures



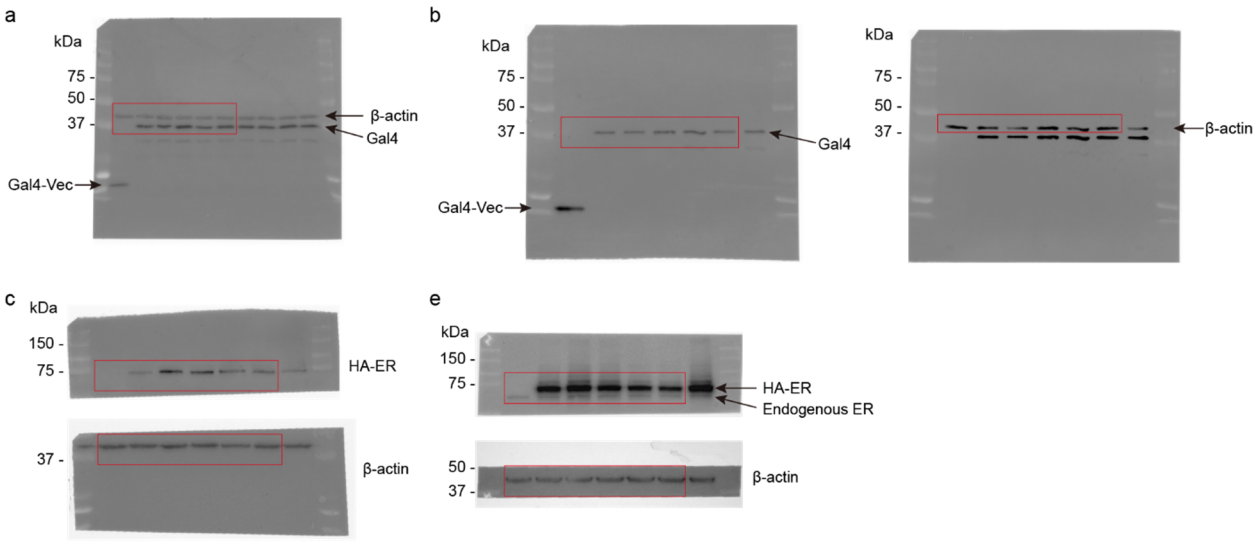
Supplementary Figure 1. Uncropped gel source images used in Extended Data Fig. 4.



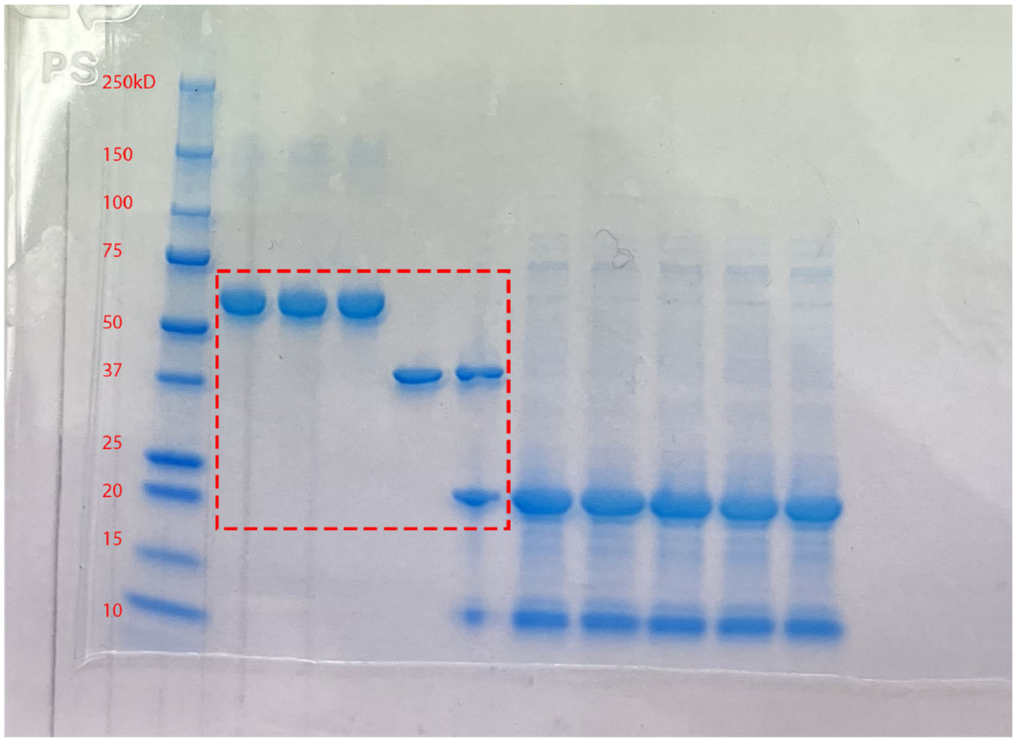
Supplementary Figure 2. Uncropped gel source images used in Extended Data Fig. 7.



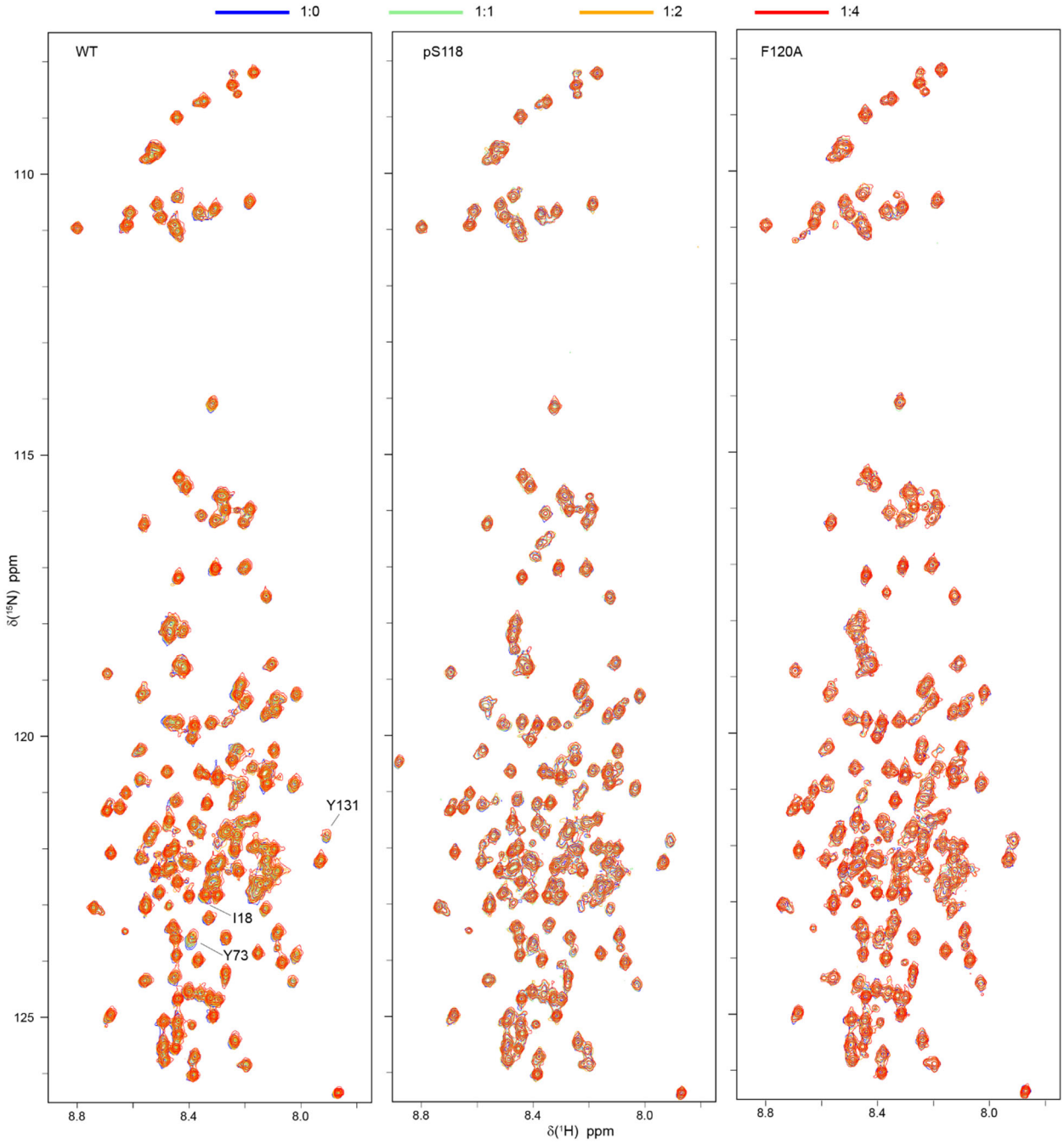
Supplementary Figure 3. Uncropped gel source images used in Extended Data Fig. 8.



Supplementary Figure 4. Uncropped gel source images used in Fig. 4.



Supplementary Figure 5. Uncropped gel source image used in Extended Data Fig. 9.



Supplementary Figure 6. ^{1}H-^{15}N HSQC spectra of ER-NTD variants with TIF2-QRD titration. Spectra were acquired for three ER-NTD variants (WT, pS118, and F120A) at 30 μM concentration. Each variant was measured at ER-NTD:TIF2-QRD molar ratios of 1:0, 1:1, 1:2, and 1:4.

Supplementary Table 1. Characteristics of 48 nuclear receptor NTDs

N: Chain length (number of amino acid residues)
FCR: Fraction of charged residues
SHP: Sequence hydrophobic patterning (normalized)
$\nu_{simulated}$: Flory exponent from coarse-grained simulations

Receptor	*N*	FCR	SHP	$\nu_{simulated}$
0B/DAX1/NR0B1	204	0.20	0.64	0.50
0B/SHP/NR0B2	15	0	0.31	0.51
1A/TRα/THRA	52	0.37	0.41	0.53
1A/TRβ/THRB	106	0.35	0.52	0.54
1B/RARα/RARA	87	0.08	0.61	0.53
1B/RARβ/RARB	87	0.13	0.61	0.53
1B/RARγ/RARG	89	0.16	0.58	0.53
1C/PPARα/PPARA	101	0.17	0.58	0.56
1C/PPARβ/PPARD	73	0.29	0.49	0.57
1C/PPARγ/PPARG	138	0.26	0.62	0.56
1D/REV-ERBα/NR1D1	131	0.10	0.59	0.52
1D/REV-ERBβ/NR1D2	102	0.17	0.54	0.54
1F/RORα/RORA	72	0.24	0.48	0.54
1F/RORβ/RORB	20	0.25	0.34	0.55
1F/RORγ/RORC	30	0.37	0.35	0.55
1H/FXRα/NR1H4	136	0.18	0.62	0.52
1H/LXRα/NR1H3	97	0.24	0.55	0.54
1H/LXRβ/NR1H2	86	0.27	0.55	0.55
1I/VDR/VDR	23	0.26	0.35	0.54
1I/PXR/PXR	40	0.35	0.41	0.56
1I/CAR/NR1I3	10	0.50	0.18	0.52
2A/HNF4α/HNF4A	59	0.17	0.50	0.55
2A/HNF4γ/HNF4G	11	0.09	0.24	0.51
2B/RXRα/RXRA	134	0.11	0.65	0.53
2B/RXRβ/RXRB	204	0.15	0.69	0.52
2B/RXRγ/RXRG	138	0.10	0.65	0.53
2C/TR2/NR2C1	112	0.21	0.58	0.55
2C/TR4/NR2C2	116	0.17	0.58	0.54
2E/TLX/NR2E1	15	0.2	0.31	0.52
2E/PNR/NR2E3	46	0.17	0.45	0.54
2F/COUP-TFα/NR2F1	85	0.13	0.52	0.54
2F/COUP-TFβ/NR2F2	78	0.12	0.54	0.54

Receptor	*N*	FCR	SHP	$v_{simulated}$
2F/COUP-TFγ/NR2F6	55	0.25	0.46	0.55
3A/ERα/ESR1	184	0.18	0.65	0.51
3A/ERβ/ESR2	148	0.20	0.61	0.52
3B/ERRα/ESRRA	78	0.26	0.54	0.55
3B/ERRβ/ESRRB	102	0.22	0.55	0.53
3B/ERRγ/ESRRG	127	0.24	0.58	0.54
3C/AR/AR	558	0.18	0.79	0.50
3C/GR/NR3C1	420	0.21	0.74	0.53
3C/MR/NR3C2	602	0.18	0.78	0.51
3C/PR/PGR	566	0.18	0.82	0.53
4A/NGF1-B/NR4A1	266	0.13	0.74	0.52
4A/NURR1/NR4A2	262	0.17	0.71	0.51
4A/NOR-1/NR4A3	291	0.18	0.75	0.52
5A/SF-1/NR5A1	12	0.50	0.25	0.56
5A/LRH-1/NR5A2	85	0.27	0.52	0.53
6A/GCNF/NR6A1	59	0.25	0.48	0.55